



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/751,574	12/29/2000	Cary Lee Bates	IBM/ 166	8366
26517	7590	11/07/2005	EXAMINER	
WOOD, HERRON & EVANS, L.L.P. (IBM)			HUYNH, CONG LACT	
2700 CAREW TOWER			ART UNIT	
441 VINE STREET			PAPER NUMBER	
CINCINNATI, OH 45202			2178	
DATE MAILED: 11/07/2005				

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	09/751,574	BATES ET AL.
	Examiner	Art Unit
	Cong-Lac Huynh	2178

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 11 August 2005.
- 2a) This action is **FINAL**. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1-43 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) Claim(s) 40 and 41 is/are allowed.
- 6) Claim(s) 1-39, 42 and 43 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- 1) Notice of References Cited (PTO-892)
- 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
- 3) Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
 Paper No(s)/Mail Date _____.
- 4) Interview Summary (PTO-413)
 Paper No(s)/Mail Date. _____.
- 5) Notice of Informal Patent Application (PTO-152)
- 6) Other: _____.

DETAILED ACTION

1. This action is responsive to communications: amendment filed 8/11/05 to the application filed on 12/29/00.
2. Claims 1-43 are pending in the case. Claims 1, 23-24, 37, 39-40, 42 are independent claims.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

4. Claims 1, 9, 12, 24, 29-30, 37-38 remain rejected under 35 U.S.C. 102(e) as being anticipated by Peterman (US Pat No. 6,480,838 B1, 11/12/02, filed 6/9/00, priority 4/1/98).

Regarding independent claim 1, Peterman discloses:

- scanning a plurality of documents for variants of a linguistic term (col 2, lines 40-58 and col 3, lines 51-55: "... *the mechanism of the invention determines variant words associated with the received search request ...A database with OCR scanned documents is then searched for the variant words*"; col 3, lines 3-7: word variant searching of electronic documents using optical character recognition)

- tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents during scanning to determine an acceptable usage of the linguistic term, wherein each variant of the linguistic term found in each document is of the type that is intentionally chosen by an author of such document (col 2, lines 52-58 and col 6, lines 25-43: determining the *possible variants of the search word* that has a probability of *occurrence* higher than a designated predetermined threshold probability *via the OCR program* inherently shows tracking the occurrences of the word variant *during scanning* to select the word variant whose occurrences *higher than a predetermined threshold* as an acceptable usage of the linguistic term; since the words in each document are used by the document's author, each variant of the linguistic term in the document is clearly chosen by the document's author)

Regarding claim 9, which is dependent on claim 1, Peterman discloses that the linguistic term comprises a single word (col 2, lines 52-58).

Regarding claim 12, which is dependent on claim 1, Peterman discloses that the plurality of variants differ from one another based upon at least one of punctuation, spelling, capitalization, hyphenation, and definition (figure 2: the alternatives for each character and combination of characters show that the variants differ from one another based upon spelling; col 10, lines 39-47).

Claims 24, 29, 30 are an apparatus of method claims 1, 9, 12, and are rejected under the same rationale.

Claims 37-38 are a program product of method claim 1, and are rejected under the same rationale.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

7. Claims 2-8, 10, 13-14, 16-18, 20, 25-29, 31-33, 35 remain rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman as applied to claim 1 above, and further in view of Anick et al. (US Pat No. 6,519,586 B2, 2/11/03, filed 8/6/99).

Regarding claim 2, which is dependent on claim 1, Peterman does not disclose retrieving the plurality of documents from a network, wherein scanning the plurality of documents includes scanning each document subsequent to retrieval of the document from the network.

Anick discloses that retrieving the plurality of documents from a network, wherein scanning the plurality of documents includes scanning each document subsequent to retrieval of the document from the network (col 3, lines 51-67, col 4, line 45 to col 5, line 25, col 6, lines 34-65, and figure 3A: retrieving documents from the Internet via submitting queries, then the retrieved document is searched for different terms of a word that can be used to identify a key concept or facet where searching through a document in such a way is equivalent to scanning the document).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman since Anick discloses retrieving documents from a network where scanning documents is subsequent to retrieving providing the advantage to expand the scope of searching documents within a local database as in Peterman into searching documents over the Internet.

Regarding claim 3, which is dependent on claim 2, Peterman does not disclose retrieving the plurality of documents from at least one Internet web site in response to a user browsing the at least one Internet web site, and wherein scanning the plurality of documents includes scanning each document upon retrieval of that document from the at least one Internet web site.

Anick discloses retrieving the plurality of documents from at least one Internet web site in response to a user browsing the at least one Internet web site, and wherein scanning the plurality of documents includes scanning each document upon retrieval of that document from the at least one Internet web site (col 3, lines 51-67, figure 3A: retrieving documents from a Internet web site in response to user entering the topic, then the searching (equivalent to scanning as mentioned in claim 2) is performed on the retrieved document from said web site).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman since Anick discloses retrieving documents from a network where scanning documents is subsequent to retrieving providing the advantage to expand the scope of searching documents within a local database as in Peterman into searching documents over the Internet.

Regarding claim 4, which is dependent on claim 2, Peterman and Anick do not disclose determining whether a retrieved document has already been scanned before scanning the retrieved document.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Peterman and Anick to incorporate determining whether a retrieved document has already been scanned before scanning the retrieved document for the following reason. It was well known that each computer system has a cache for storing the retrieved documents as well as a cookie file for storing user actions performed while in the Internet and for preventing a repeated action if it is

Art Unit: 2178

already in cache. Therefore, the system will check to see if a document is scanned yet via the cache and cookie file before performing scanning documents.

Regarding claim 5, which is dependent on claim 2, Peterman does not disclose determining whether to scan a retrieved document based upon a source parameter associated with the linguistic term.

Anick discloses determining whether to scan a retrieved document based upon a source parameter associated with the linguistic term (col 8, line 18 to col 9, line 67: a list of parameters provided to users for changing the selection of facets derived from the retrieved documents).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick to Peterman since Anick determining whether to scan a retrieved document based upon a source parameter providing the advantage of selecting different options on scanning via selecting an appropriate parameter.

Regarding claim 6, which is dependent on claim 1, Peterman does not disclose browsing a second plurality of documents retrieved from at least one Internet web site in response to user input, wherein scanning the first plurality of documents is performed concurrently with browsing the second plurality of documents.

Anick discloses:

Art Unit: 2178

- browsing a plurality of documents retrieved from at least one Internet web site in response to user input (col 3, lines 51-58; figures 3A-B: searching and browsing web pages in response to user input of topics on the Internet)
- scanning a plurality of documents on the Internet (col 5, lines 52-63)

Anick does not explicitly disclose that the above two steps occur concurrently.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include the concurrency of said two steps into Anick for the following reason. Anick discloses that browsing for a desired document on the Internet and measuring the facet frequency are occurred on a web page (figures 3A-B; col 5, lines 52-63 and col 6, lines 34-65). The fact that the web browser receives the user input on the frames of a web page for performing the two different functions suggests that the two steps can be carried out concurrently to preserve the consistence of data in the frames of the web page.

Regarding claim 7, which is dependent on claim 1, Peterman does not disclose that scanning the first plurality of documents is performed in a background thread while documents from the second plurality of documents are being browsed.

As mentioned in claim 6 above, Anick suggests that steps scanning and browsing documents be occurred concurrently for preserving the consistency of data presented in the web page. Therefore, it is further suggested that the two actions, scanning documents and browsing documents be performed parallelly in two different threads.

Regarding claim 8, which is dependent on claim 7, Peterman does not disclose that scanning the first plurality of documents includes scanning documents stored in a local history cache. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Peterman to incorporate scanning documents stored in a local history cache since it was well known that cache is a local storage of a computer with a plurality of stored documents. Therefore, scanning documents includes documents stored in cache.

Regarding claim 10, which is dependent on claim 1, Peterman does not disclose that the linguistic term comprises a phrase.

Anick discloses that the linguistic term comprises a phrase (figure 3B and col 5, lines 52-63).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for expanding the scope of considering the variant of a linguistic term as a phrase including at least two words instead of a single word as in Peterman.

Regarding claim 13, which is dependent on claim 1, Peterman does not disclose that scanning the plurality of documents includes scanning a document for an enumerated variant of the linguistic term.

Anick discloses scanning the plurality of documents includes scanning a document for an enumerated variant of the linguistic term (col 6, lines 20-33: the fact that the facet list

is retained for successive searches shows that the number of facets, which are the variants of a linguistic term, is counted and determined; in other words, the scanning is for an enumerated variant of a linguistic term).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for easily choosing one of the enumerated variants of a linguistic term.

Regarding claim 14, which is dependent on claim 1, Peterman does not disclose that scanning the plurality of documents includes scanning a document for an unenumerated variant of the linguistic term.

Anick discloses that scanning the plurality of documents includes scanning a document for an unenumerated variant of the linguistic term (col 5, lines 52-63: the fact that a user can choose to *list any desired number of facet phrases* in response to the user's click of a facet selection box for the facets corresponding to the scanned documents suggests an unenumerated variant of the linguistic term when scanning a document).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for flexibly using any variant, which is unenumerated, of a linguistic term.

Regarding claim 16, which is dependent on claim 1, Peterman does not disclose that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon locations of such occurrences within the plurality of documents.

Anick discloses tracking relative occurrences of the plurality of variants includes weighting occurrences based upon locations of such occurrences within the plurality of documents (col 10, 1-24: weighting facet occurrences on only the terms *appearing near the beginning of the documents*, those appearing in the first n document lines).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for obtaining an effective way to track the relative occurrences of the variants based on the variants' locations in the documents rather than tracking the variants on the entire documents, which is a time consuming.

Regarding claim 17, which is dependent on claim 1, Peterman does not disclose that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon document types of the documents within which such occurrences are found.

Anick discloses that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon document types of the documents within which such occurrences are found (figure 3B, col 10, lines 12-24, col 3, lines 51-67: weighting the density of the terms, which is equivalent to the term occurrences, in a document where a document can be a form of articles, papers, statements, correspondence, etc.).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for obtaining an effective way to track

the relative occurrences of the variants based on the document type rather than tracking all of the documents, which helps speeding up the tracking process.

Regarding claim 18, which is dependent on claim 1, Peterman does not disclose storing a variant of the linguistic term in an electronic dictionary.

Anick discloses storing a variant of the linguistic term in an electronic dictionary (figures 3A-B: lists of variants of linguistic terms are stored and displayed for users to select).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for the following reason. Anick discloses storing lists of variants of linguistic terms for users to select providing the advantage to apply said storing to Peterman to store the variants of linguistic terms with high occurrences for rapidly searching and retrieving documents.

Regarding claim 20, which is dependent on claim 1, Peterman does not disclose tracking relative occurrences of the plurality of variants includes storing context information associated with each occurrence of a variant of the linguistic term.

Anick discloses tracking relative occurrences of the plurality of variants includes storing context information associated with each occurrence of a variant of the linguistic term (col 6, lines 46-65: encoding strong, long-live information relating to relationships and attributes which serve well as the lexical constructs from which dispersion rates are derived suggests that said information relating to the occurrences of the variants for a particular word be stored in a database for searching or retrieving documents where the

database of words and phrases is considered as similar as an electronic dictionary for storing terms).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for the following reason. Anick discloses storing lists of variants of linguistic terms for users to select providing the advantage to apply said storing to Peterman to store the variants of linguistic terms with their associated occurrences for easily searching and retrieving desired variants.

Claims 25-28 are for an apparatus of method claims 2, 4, 5, 7, and are rejected under the same rationale.

Claims 29, 31-33, 35 are for an apparatus of method claims 10, 13-14, 16, 20 and are rejected under the same rationale.

8. Claim 11 is rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman as applied to claim 1 above, and further in view of Goiffon et al. (US Pat No. 6,453,312 B1, 9/17/02, filed 10/14/98).

Regarding claim 11, which is dependent on claim 1, Peterman does not disclose that the linguistic term comprises an acronym.

Goiffon discloses that the word elements related to stored character strings that include abbreviations, acronyms, plural formats, possessive formats and the like where the word and word variants are used to develop a query (col 4, lines 17-26).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Goiffon into Peterman since Goiffon discloses the word or word variant, which is equivalent to a linguistic term, used for query and searching, providing the advantage to incorporate into Peterman for further including an acronym as a linguistic term beside any word type in searching a term variant in a document.

9. Claim 15 remains rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman and Anick as applied to claims 1 and 14 above, and further in view of Komissarchik et al. (US Pat No. 5,799,276, 8/25/98, filed 11/7/95).

Regarding claim 15, which is dependent on claim 14, Peterman and Anick do not disclose that scanning the document for the unenumerated variant of the linguistic term includes scanning the document using phonetic comparison.

Komissarchik discloses that a single phonetic transcription of a word in a dictionary can be replaced with a *plurality of phonetic transcriptions of phonetically permissible variants of the word* (col 3, line 63 to col 4, line 15 and col 11, line 50 to col 12, line 13).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Komissarchik into Peterman and Anick since Komissarchik teaches that it is possible to replace a word in a dictionary with phonetically permissible

variants of the word thus motivating to phonetically scanning documents for the variants of a word instead of lexically scanning as in Peterman and Anick utilizing the phonetic comparison.

10. Claims 19, 21-22, 34, 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Peterman and Anick as applied to claims 1 and 18 above, and further in view of Angiulo et al. (US Pat No. 6,044,387, 3/28/00, filed 9/10/97).

Regarding claim 19, which is dependent on claim 18, Peterman and Anick do not disclose spell checking a document using the electronic dictionary subsequent to storing the variant in the electronic dictionary.

Angiulo discloses spell checking on HTML documents after the HTML documents are stored in the database (figures 1-4, col 4, line 63 to col 5, line 16).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Angiulo into Peterman and Anick for the following reason.

Angiulo discloses spell checking the HTML document using the spell checker program, where it was well known that any spell checker program uses an electronic dictionary for checking the spelling, and where these documents are stored in a database thus motivating to apply the spell checking program on a document subsequent to storing the variant in the electronic dictionary in Peterman since a whole document when being stored includes a plurality of words which can be the variant of a linguistic term.

Regarding claim 21, which is dependent on claim 1, Peterman does not disclose scanning a document for a spell definition tag that identifies a variant of the linguistic term.

Angiulo discloses spell checking a plurality of files, which are web page documents for a web site (col 2, lines 25-67) where the spell checking applies the conventional spell checking technique such as providing alternative spellings for users to select, or giving users the option to correct the spelling and add to word to the dictionary (col 1, lines 10-49).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Angiulo scanning a document for a spell definition tag identifying a variant of the linguistic term since spell checking a web document implies scanning or parsing said document to identify the incorrectly spelled word.

Further, since variant spellings of a word are provided in the process of spell checking a web document, it is suggested that identifying a variant of the linguistic term be performed.

Also, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Angiulo into Peterman since Angiulo discloses and suggests scanning a document for a spell definition to identify a variant of the linguistic term providing the advantage to incorporate the spell check feature into Peterman for rapidly and effectively providing users documents without spell errors.

Regarding claim 22, which is dependent on claim 1, Peterman does not disclose scanning the plurality of documents and tracking relative occurrences are performed responsive to detecting a variant of the linguistic term during spell checking of a document.

Instead, Peterman does teach scanning the plurality of documents and tracking relative occurrences of the variant of a linguistic term where tracking occurrences of the variant of a linguistic term inherently shows detecting a variant of the linguistic term (col 2, lines 40-58, col 3, lines 51-55, col 6, lines 25-43).

Angiulo discloses scanning a document for spell checking of the document (col 4, line 63 to col 5, line 16).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Angiulo into Peterman to incorporate the spell checking program of Mueller into scanning documents of Peterman for detecting the variant of a linguistic term since scanning a document in Peterman is for looking through the document for detecting the variant and scanning a document in Angiulo is also for looking through the document for detecting spell errors.

Claims 34 and 36 are for an apparatus of method claims 19 and 21, and are rejected under the same rationale.

11. Claims 23 and 39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anick et al. (US Pat No. 6,519,586, 2/11/03, filed 8/6/99) in view of Peterman (US Pat No. 6,480,838 B1, 11/12/02, filed 6/9/00).

Regarding independent claim 23, Anick discloses:

- browsing a plurality of web sites on the Internet in response to user input (col 3, lines 51-58; figures 3A-B: searching and browsing web pages in response to user input of topics on the Internet)

Anick does not disclose:

- tracking relative occurrences of a plurality of variants of a linguistic term found in the plurality of web sites to determine an acceptable usage of the linguistic term (col 5, lines 52-63: measuring the frequency of facet's occurrence for a retrieved document)

Peterman discloses:

- tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents during scanning to determine an acceptable usage of the linguistic term, wherein each variant of the linguistic term found in each web site is of the type that is intentionally chosen by an author of such web site (col 2, lines 52-58 and col 6, lines 25-43: determining the *possible variants of the search word* that has a probability of *occurrence* higher than a designated predetermined threshold probability *via the OCR program* inherently shows tracking the occurrences of the word variant *during scanning* to select the word

variant whose occurrences *higher than a predetermined threshold* as an acceptable usage of the linguistic term; since the words in each web site are used by the website's author, each variant of the linguistic term in the website is clearly chosen by the website's author)

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Peterman into Anick since Peterman discloses tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents to determine an acceptable usage of the linguistic term. This feature provides the advantage to incorporate into Anick for tracking relative occurrences of the linguistic term found in the plurality of the web documents browsed in the Internet to determine an acceptable usage of the linguistic term since in term of the contents of documents, the documents in Anick are the same as the documents in Peterman.

Anick and Peterman do not disclose that the two steps browsing and tracking occur concurrently.

Anick discloses that the invention is implemented in an interface to a document search engine facilitating document searching and browsing on the World Wide Web (col 3, lines 51-61).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents to determine an acceptable usage of the linguistic term since said tracking is a type of searching, and the fact that the document search engine for searching and browsing in an interface

implies that searching and browsing occur concurrently. The browsing and tracking, by analogy, can be performed concurrently in an interface the same way as browsing and searching do.

Claim 39 is for a program product of method claim 23, and is rejected under the same rationale.

12. Claims 42-43 are rejected under 35 U.S.C. 103(a) as being unpatentable over Angiulo et al. (US Pat No. 6,044,387, 3/28/00, filed 9/10/97).

Regarding independent claim 42, Angiulo discloses:

- detecting a tag within a document retrieved from the Internet that identifies an acceptable variant of a linguistic term for carrying out spell checking (col 2, lines 25-67, col 4, line 36 to col 5, line 16, col 1, lines 11-49: spell checking the web documents shows that the document is retrieved from the Internet where a web document includes tags where the terms that do not constitute a properly spelled word are identified and prompted to users with several suggested replacements or a user can enter a word to replace the incorrect word)
- in response to detecting said tag, automatically adding the acceptable variant of the linguistic term to an electronic dictionary (col 6, lines 40-63)

Angiulo does not disclose explicitly such a spell definition tag. However, it was well known that to include different types of tags for instructing a required command

including a spell definition tag in a web document. Therefore, it would have been obvious at the time of the invention was made to combine such tag in a web document for performing spell checking to detect a variant for a linguistic term in the web document. The combination of such a tag into Angiulo would facilitate the spell checking process based on the spell definition defined in the tags.

Regarding claim 43, Angiulo discloses that detecting the spell definition tag is performed during user browsing of the Internet (col 2, lines 25-67: spell checking the Internet pages implies that the detecting is performed during user browsing of the Internet).

13. Claims 1-10, 12-18, 20, 24-33, 35, 37-38 remain rejected under 35 U.S.C. 103(a) as being unpatentable over Anick et al. (US Pat No. 6,519,586 B2, 2/11/03, filed 8/6/99).

Regarding independent claim 1, Anick discloses:

- searching a plurality of documents for variants of a linguistic term (figures 3A-B: searching for variants of a linguistic term in the retrieved documents)
- tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents during searching to determine an acceptable usage of the linguistic term (col 5, lines 52-63: measuring the frequency of facet's occurrences in the documents)

Anick does not use the word "scanning document" for variants of a linguistic term.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to have scanning documents for variant of a linguistic term since searching said variant in the documents means looking through the documents to find out the word variant. That is scanning the documents for the word variant.

Regarding claim 2, which is dependent on claim 1, Anick discloses that retrieving the plurality of documents from a network, wherein scanning the plurality of documents includes scanning each document subsequent to retrieval of the document from the network (col 3, lines 51-67, col 4, line 45 to col 5, line 25, col 6, lines 34-65, and figure 3A: retrieving documents from the Internet via submitting queries, then the retrieved document is searched for different terms of a word that can be used to identify a key concept or facet where searching through a document in such a way is equivalent to scanning the document).

Regarding claim 3, which is dependent on claim 2, Anick discloses retrieving the plurality of documents from at least one Internet web site in response to a user browsing the at least one Internet web site, and wherein scanning the plurality of documents includes scanning each document upon retrieval of that document from the at least one Internet web site (col 3, lines 51-67, figure 3A: retrieving documents from a Internet web site in response to user entering the topic, then the searching (equivalent to scanning as mentioned in claim 2) is performed on the retrieved document from said web site).

Regarding claim 4, which is dependent on claim 2, Anick does not disclose determining whether a retrieved document has already been scanned before scanning the retrieved document.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to incorporate determining whether a retrieved document has already been scanned before scanning the retrieved document for the following reason. It was well known that each computer system has a cache for storing the retrieved documents as well as a cookie file for storing user actions performed while in the Internet and for preventing a repeated action if it is already in cache. Therefore, the system will check to see if a document is scanned yet via the cache and cookie file before performing scanning documents.

Regarding claim 5, which is dependent on claim 2, Anick discloses determining whether to scan a retrieved document based upon a source parameter associated with the linguistic term (col 8, line 18 to col 9, line 67: a list of parameters provided to users for changing the selection of facets derived from the retrieved documents).

Regarding claim 6, which is dependent on claim 1, Anick discloses:

- browsing a plurality of documents retrieved from at least one Internet web site in response to user input (col 3, lines 51-58; figures 3A-B: searching and browsing web pages in response to user input of topics on the Internet)

- scanning a plurality of documents on the Internet (col 5, lines 52-63)

Anick does not explicitly disclose that the above two steps occur concurrently.

However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include the concurrency of said two steps into Anick for the following reason. Anick discloses that browsing for a desired document on the Internet and measuring the facet frequency are occurred on a web page (figures 3A-B; col 5, lines 52-63 and col 6, lines 34-65). The fact that the web browser receives the user input on the frames of a web page for performing the two different functions suggests that the two steps can be carried out concurrently to preserve the consistence of data in the frames of the web page.

Regarding claim 7, which is dependent on claim 1, as mentioned in claim 6 above, Anick suggests that steps scanning and browsing documents be occurred concurrently for preserving the consistency of data presented in the web page. Therefore, it is further suggested that the two actions, scanning documents and browsing documents be performed parallelly in two different threads.

Regarding claim 8, which is dependent on claim 7, Anick does not disclose that scanning the first plurality of documents includes scanning documents stored in a local history cache. However, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Peterman to incorporate scanning documents stored in a local history cache since it was well known that cache is a local

storage of a computer with a plurality of stored documents. Therefore, scanning documents includes documents stored in cache.

Regarding claim 9, which is dependent on claim 1, Anick does not disclose that the linguistic term comprises a word. Instead, Anick discloses that the linguistic term comprises a phrase (figure 3B and col 5, lines 52-63).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Anick to include that the linguistic term comprises a word since the phrase in Anick includes at least a word.

Regarding claim 10, which is dependent on claim 1, Anick discloses that the linguistic term comprises a phrase (figure 3B and col 5, lines 52-63).

Regarding claim 12, which is dependent on claim 1, Anick discloses that the plurality of variants differ from one another based upon at least one of punctuation, spelling, capitalization, hyphenation, and definition (figure 3B: variants in 325 differ from definition).

Regarding claim 13, which is dependent on claim 1, Anick discloses scanning the plurality of documents includes scanning a document for an enumerated variant of the linguistic term (col 6, lines 20-33: the fact that the facet list is retained for successive searches shows that the number of facets, which are the variants of a linguistic term, is

counted and determined; in other words, the scanning is for an enumerated variant of a linguistic term).

Regarding claim 14, which is dependent on claim 1, Anick discloses that scanning the plurality of documents includes scanning a document for an unenumerated variant of the linguistic term (col 5, lines 52-63: the fact that a user can choose *to list any desired number of facet phrases* in response to the user's click of a facet selection box for the facets corresponding to the scanned documents suggests an unenumerated variant of the linguistic term when scanning a document).

Regarding claim 16, which is dependent on claim 1, Anick discloses tracking relative occurrences of the plurality of variants includes weighting occurrences based upon locations of such occurrences within the plurality of documents (col 10, 1-24: weighting facet occurrences on only the terms *appearing near the beginning of the documents*, those appearing in the first n document lines).

Regarding claim 17, which is dependent on claim 1, Anick discloses that tracking relative occurrences of the plurality of variants includes weighting occurrences based upon document types of the documents within which such occurrences are found (figure 3B, col 10, lines 12-24, col 3, lines 51-67: weighting the density of the terms, which is equivalent to the term occurrences, in a document where a document can be a form of articles, papers, statements, correspondence, etc.).

Regarding claim 18, which is dependent on claim 1, Anick discloses storing a variant of the linguistic term in an electronic dictionary (figures 3A-B: lists of variants of linguistic terms are stored and displayed for users to select).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Anick into Peterman for the following reason. Anick discloses storing lists of variants of linguistic terms for users to select providing the advantage to apply said storing to Peterman to store the variants of linguistic terms with high occurrences for rapidly searching and retrieving documents.

Regarding claim 20, which is dependent on claim 1, Anick discloses tracking relative occurrences of the plurality of variants includes storing context information associated with each occurrence of a variant of the linguistic term (col 6, lines 46-65: encoding strong, long-live information relating to relationships and attributes which serve well as the lexical constructs from which dispersion rates are derived suggests that said information relating to the occurrences of the variants for a particular word be stored in a database for searching or retrieving documents where the database of words and phrases is considered as similar as an electronic dictionary for storing terms).

Claims 24-29 are for an apparatus of method claims 1, 2, 4, 5, 7, and are rejected under the same rationale.

Claims 30-33, 35 are for an apparatus of method claims 12-14, 16, 20 and are rejected under the same rationale.

Claims 37-38 are for a program product of method claim 1, and are rejected under the same rationale.

14. Claim 11 is rejected under 35 U.S.C. 103(a) as being unpatentable over Anick as applied to claim 1 above, and further in view of Goiffon et al. (US Pat No. 6,453,312 B1, 9/17/02, filed 10/14/98).

Regarding claim 11, which is dependent on claim 1, Anick does not disclose that the linguistic term comprises an acronym.

Goiffon discloses that the word elements related to stored character strings that include abbreviations, acronyms, plural formats, possessive formats and the like where the word and word variants are used to develop a query (col 4, lines 17-26).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Goiffon into Anick since Goiffon discloses the word or word variant, which is equivalent to a linguistic term, used for query and searching, providing the advantage to incorporate into Anick for further including an acronym as a linguistic term beside any word type in searching a term variant in a document.

15. Claim 15 is rejected under 35 U.S.C. 103(a) as being unpatentable over Anick as applied to claims 1 and 14 above, and further in view of Komissarchik et al. (US Pat No. 5,799,276, 8/25/98, filed 11/7/95).

Regarding claim 15, which is dependent on claim 14, Anick does not disclose that scanning the document for the unenumerated variant of the linguistic term includes scanning the document using phonetic comparison.

Komissarchik discloses that a single phonetic transcription of a word in a dictionary can be replaced with *a plurality of phonetic transactions of phonetically permissible variants of the word* (col 3, line 63 to col 4, line 15 and col 11, line 50 to col 12, line 13).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Komissarchik into Anick since Komissarchik teaches that it is possible to replace a word in a dictionary with phonetically permissible variants of the word thus motivating to phonetically scanning documents for the variants of a word instead of lexically scanning as in Anick, utilizing the phonetic comparison.

16. Claims 19, 21-22, 34, 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Anick as applied to claims 1 and 18 above, and further in view of Angiulo et al. (US Pat No. 6,044,387, 3/28/00, filed 9/10/97).

Regarding claim 19, which is dependent on claim 18, Anick does not disclose spell checking a document using the electronic dictionary subsequent to storing the variant in the electronic dictionary.

Angiulo discloses spell checking on HTML documents after the HTML documents are stored in the database (figures 1-4, col 4, line 63 to col 5, line 16).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Angiulo into Anick for the following reason. Angiulo discloses spell checking the HTML document using the spell checker program, where it was well known that any spell checker program uses an electronic dictionary for checking the spelling, and where these documents are stored in a database thus motivating to apply the spell checking program on a document subsequent to storing the variant in the electronic dictionary since a whole document when being stored includes a plurality of words which can be the variant of a linguistic term.

Regarding claim 21, which is dependent on claim 1, Anick does not disclose scanning a document for a spell definition tag that identifies a variant of the linguistic term.

Angiulo discloses spell checking a plurality of files, which are web page documents for a web site (col 2, lines 25-67) where the spell checking applies the conventional spell checking technique such as providing alternative spellings for users to select, or giving users the option to correct the spelling and add to word to the dictionary (col 1, lines 10-49).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have modified Angiulo scanning a document for a spell definition tag identifying a variant of the linguistic term since spell checking a web document implies scanning or parsing said document to identify the incorrectly spelled word.

Further, since variant spellings of a word are provided in the process of spell checking a web document, it is suggested that identifying a variant of the linguistic term be performed.

Also, it would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Angiulo into Anick since Angiulo discloses and suggests scanning a document for a spell definition to identify a variant of the linguistic term providing the advantage to incorporate the spell check feature into Anick for rapidly and effectively providing users documents without spell errors.

Regarding claim 22, which is dependent on claim 1, Anick and Angiulo do not disclose scanning the plurality of documents and tracking relative occurrences are performed responsive to detecting a variant of the linguistic term during spell checking of a document.

Instead, Anick does teach scanning the plurality of documents and tracking relative occurrences of the variant of a linguistic term where tracking occurrences of the variant of a linguistic term inherently shows detecting a variant of the linguistic term (figures 3A-B and col 5, lines 52-63).

Angiulo discloses scanning a document for spell checking of the document (col 4, line 64 to col 5, line 16).

It would have been obvious to one of ordinary skill in the art at the time of the invention was made to have combined Angiulo into Anick to incorporate the spell checking program of Mueller into scanning documents of Anick for detecting the variant of a linguistic term since scanning a document in Anick is for looking through the document for detecting the variant and scanning a document in Angiulo is also for looking through the document for detecting spell errors.

Claims 34 and 36 are for an apparatus of method claims 19 and 21, and are rejected under the same rationale.

Response to Arguments

17. Applicant's arguments filed 8/11/05 have been fully considered but they are not persuasive.

Applicants argue that Peterman does not disclose that each variant of the linguistic term found in each document is of the type that is intentionally chosen by an author of such document (Remarks, page 11).

Examiner respectfully disagrees.

Though Peterman does not *explicitly* disclose that each variant of the linguistic term found in each document is of the type that is intentionally chosen by an author of such document, the fact that the words in each document are used by the document's author

when writing the document implies that the variants of the linguistic term, which are words in the document, must be chosen by the document's author.

Applicants also argue that Peterman does not disclose tracking of variant due to any chosen usages of variants by authors of documents since said tracking in Peterman is due to OCR scanning errors (Remarks, page 12).

Examiner respectfully disagrees.

In Peterman, a search word is a linguistic term. Determining the possible variants of a search word based on the probability of occurrence higher than a designated predetermined threshold probability (col 2, lines 52-58 and col 6, lines 25-43) shows that a term variant that has the probability of occurrence *higher than a designated predetermined threshold probability* is considered an chosen usage of said term since only the term variant qualifies this condition is considered chosen.

Applicants claim "tracking relative occurrences of a plurality of variants of the linguistic term found in the plurality of documents during scanning..." Peterman discloses that feature "*A Database with OCR scanned documents is then searched for the variant words... variant words are determined by determining word segments that represent OCR interpretation ...of the cumulative probability for a word segment is below the threshold, the word segment is rejected as a variant word... a tree data structure is created, having branch nodes ...each branch node represents a possible delineation of a character during OCR processing*" (col 2, lines 40-67). It is clear that in Peterman, the

occurrences of the variants are found during the scanning process. The words used in the documents are of course chosen by the authors of the documents.

Regarding claim 23, Applicants argue that Anick does not disclose or suggest “tracking relative occurrences of a plurality of variants of a linguistic term and determining an acceptable usage of the linguistic term ...” and Peterman, in combination with Anick, fails to disclose that features (Remarks, page 13).

Examiner respectfully disagrees.

Peterman discloses the argued feature (see claim 1 rejection).

Regarding claims 24 and 37, these claims are amended to recite the feature that the variants of the linguistic term is the type that are intentionally chosen by authors of scanned documents (Remarks, page 15).

Examiner respectfully disagrees.

Peterman implies the argued feature. See claim 1 rejection.

Regarding claim 39, Applicants argue that the claim, which is the program of method claim 23, recites a tag in a document that identifies “an acceptable variant of a linguistic term and a definition of the linguistic term”, a feature that is not recited in claim 23 (Remarks, page 16).

Examiner respectfully disagrees.

Since it was well known that a web document is written by a markup language using tags for describing how to present the content of the document, the program of the method 23 should include tags for identifying an acceptable variant of a linguistic term and a definition of the linguistic term to reflect the feature tracking relative occurrences of a plurality of variants of a linguistic term to determine an acceptable usage of the linguistic term recited in claim 23.

Regarding claim 42, Applicants argue that the rejection of the claim is based on Angiulo combined with Anick, which is not a proper combination (Remarks, page 18).

Examiner respectfully disagrees.

Claims 42 and its dependent 43 are rejected by Angiulo alone (see the claim rejection, pages 21-22). However, the claim rejections were misplaced under the title of rejections by Angiulo in view of Anick. That error is corrected in this office action.

Allowable subject matter

18. Claims 40-41 are allowed.
19. The following is an examiner's statement of reasons for allowance: automatically scanning a plurality of documents from the Internet to identify at least one acceptable image of the term in response to determining during the comparison that a term from a document is not in the electronic dictionary, was not disclosed by, or would not have been obvious to an ordinary skill in the art over the prior art of record.

Conclusion

20. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

21. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Light (US Pat No. 6,480,835 B1, 11/12/02, filed 12/31/98).

Jacquemin et al., *Retrieving Terms and Their Variants in a Lexicalized Unification-Based Framework*, ACM August 1994, pages 132-144.

Jacquemin , Syntagmatic and Paradigmatic Representations of Term Variation, ACM June 1999, pages 341-348.

22. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Cong-Lac Huynh whose telephone number is 571-272-4125. The examiner can normally be reached on Mon-Fri (8:30-6:00).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Stephen Hong can be reached on 571-272-4124. The fax phone number for the organization where this application or proceeding is assigned is 571-273-4125.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).



Cong-Lac Huynh
Examiner
Art Unit 2178
10/28/05